Distributed QoS Control

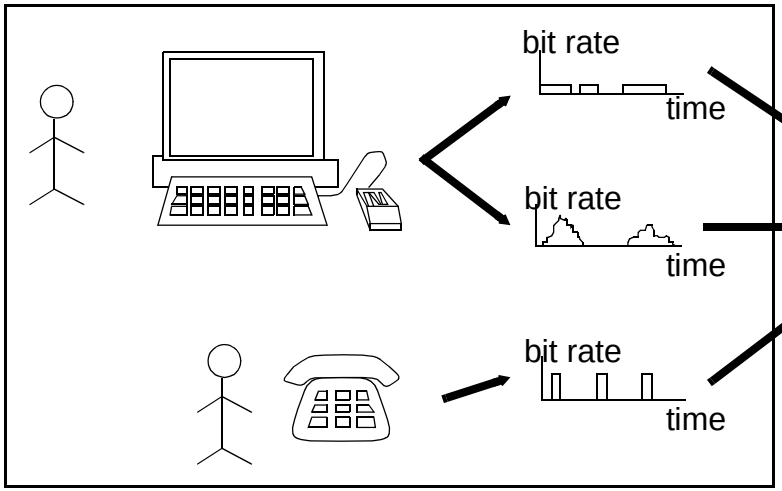
Scott Jordan
University of California, Irvine

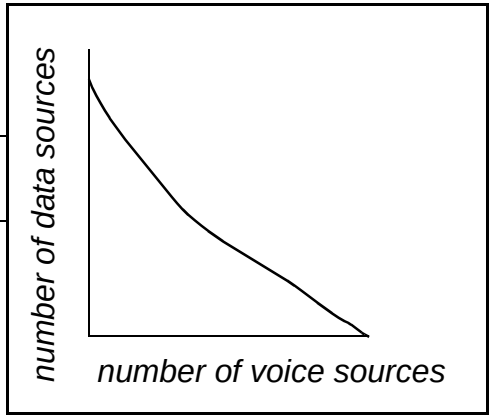
New Ideas

- *Build connections between characterizations of traffic flows, QoS requests, and network resource availability*
- *Negotiations between network and user agents regarding QoS*
- *Minimize information exchange using price & demand*

Impact

- *Reservation of network resources for each traffic flow or aggregates of flows in integrated service architectures*
- *Priority marking of packets in differentiated service architectures*
- *Automate resource management and QoS management tasks*





User

Traffic Characterization

Quality of Service (QoS)

Network

Scheduling / CAC / Policing

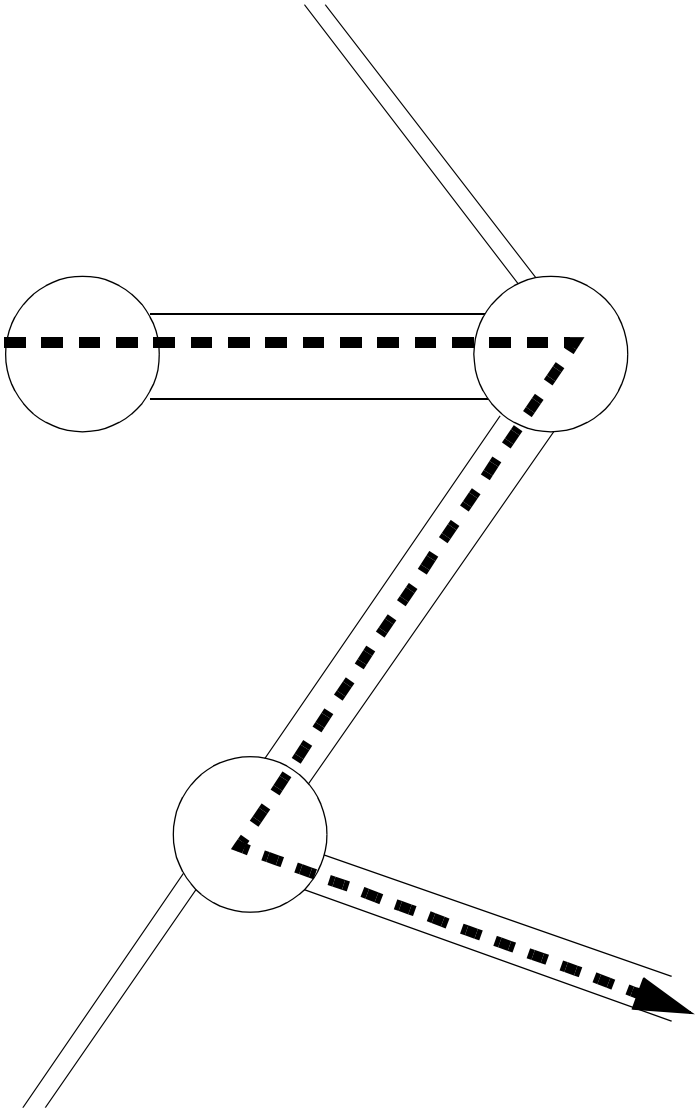
Network Capacity Characterization

Contract

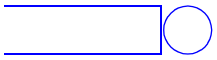
Pricing

Integrated Services

Internet - “RSVP”
ATM - virtual circuits

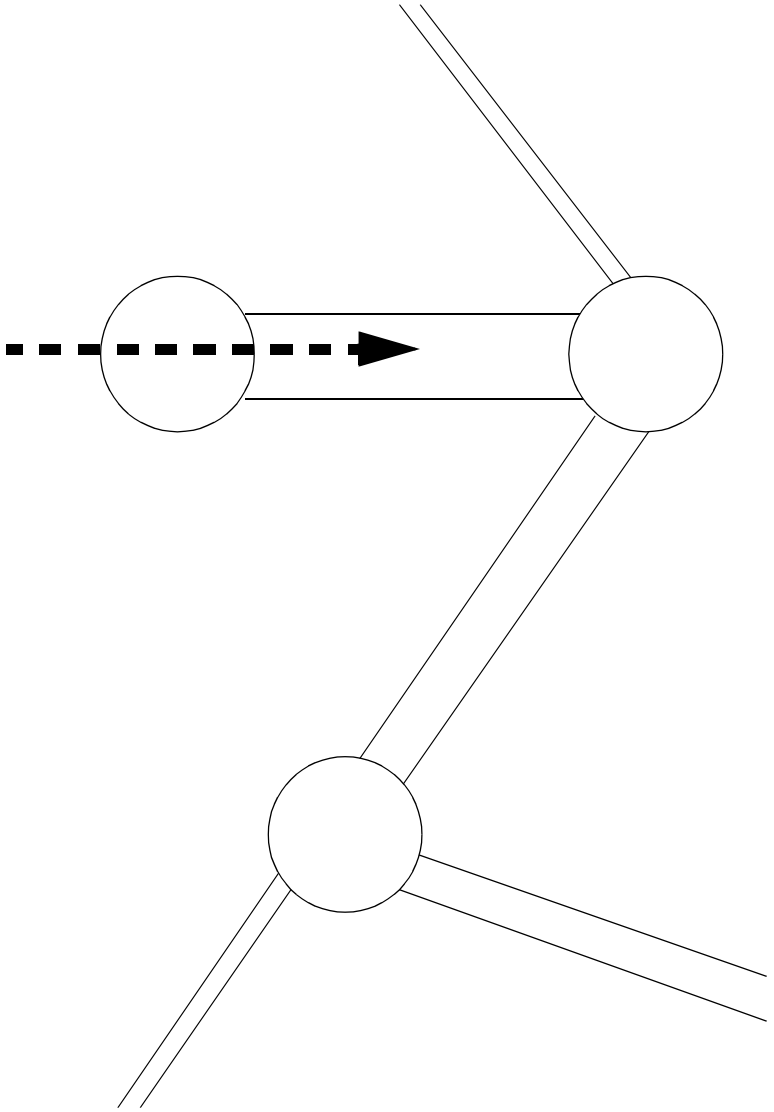


reserve buffer and bandwidth along route

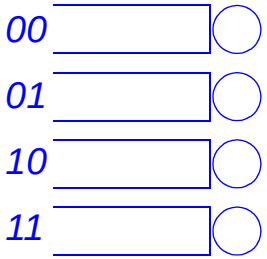


Differentiated Services

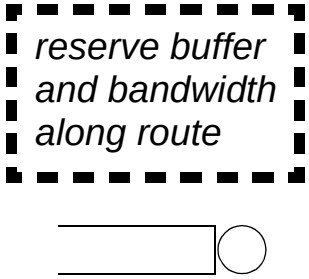
Internet - “diffServ”



label packets with traffic class



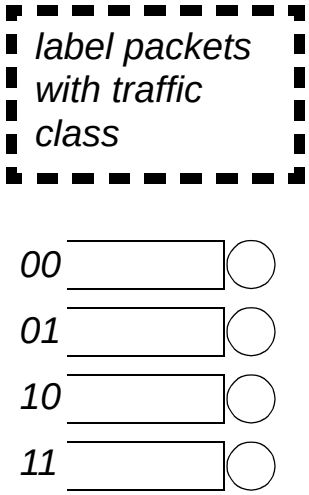
Integrated Services



How much buffer and bandwidth should I reserve ??

Depends on desired QoS and congestion !!

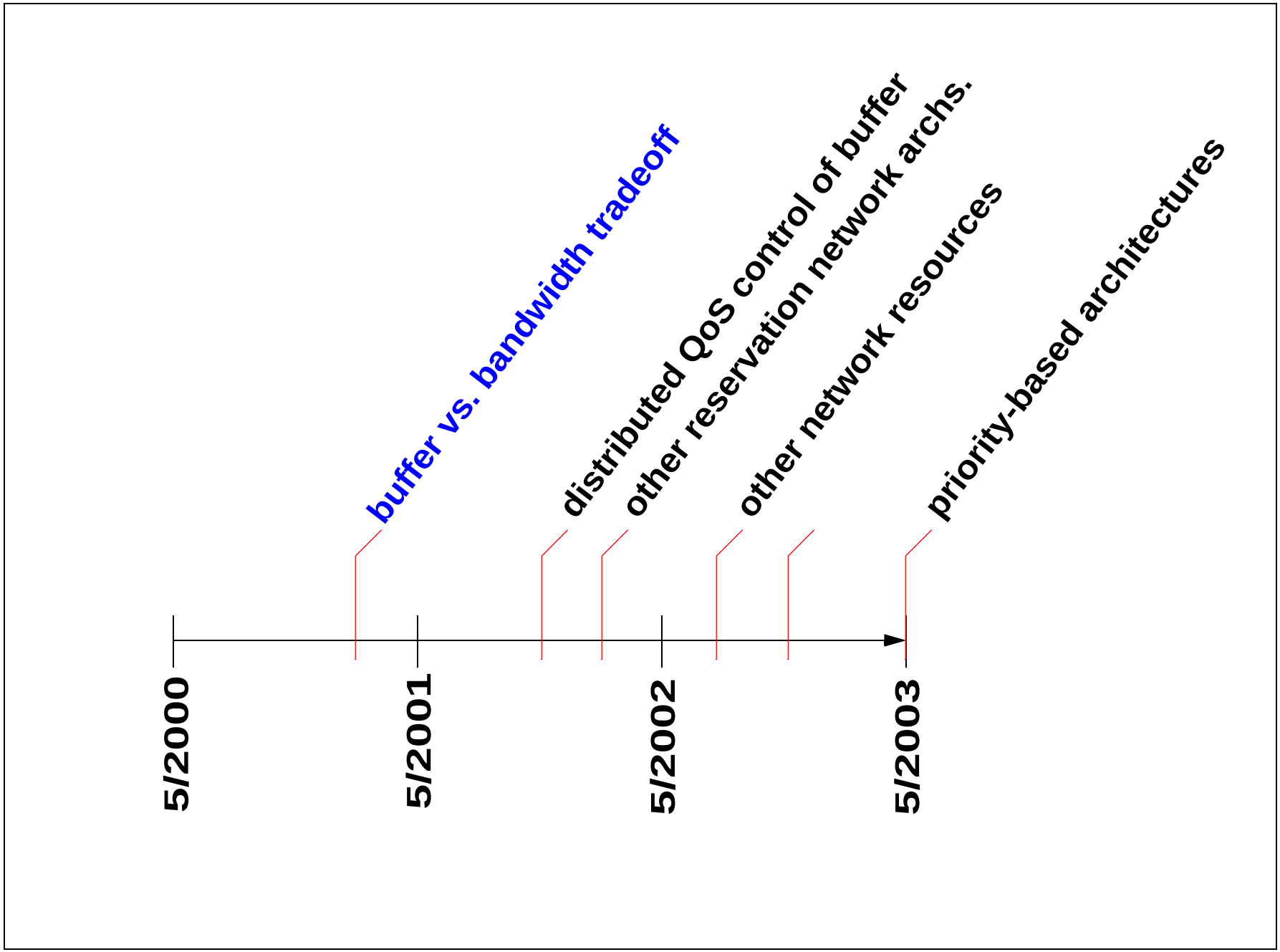
Differentiated Services



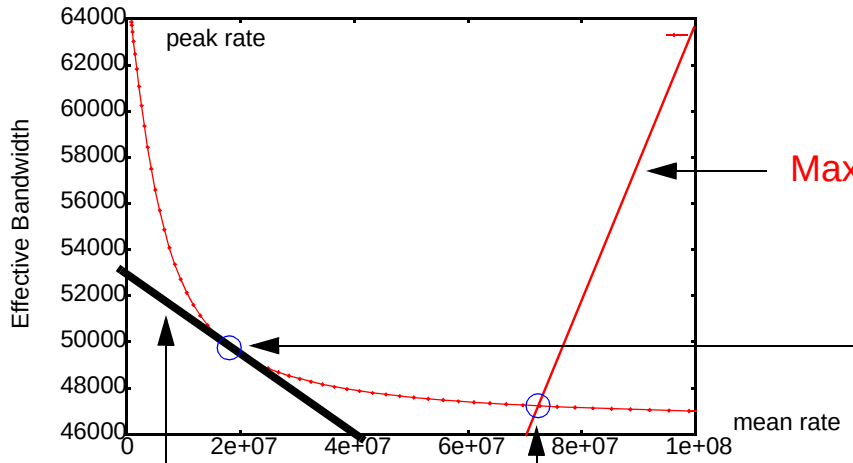
Which traffic class(es) should I use ??

Depends on desired QoS and congestion !!

Timeline



Current Research -- Optimal allocation of buffer vs. bandwidth



Maximum delay constraint

Optimal buffer and bandwidth allocation

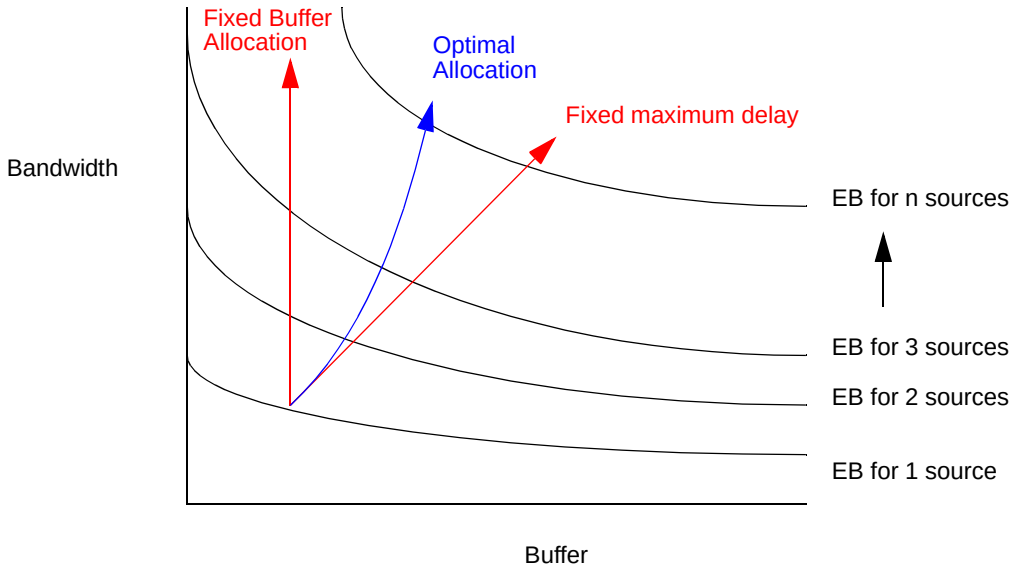
Delay-constrained allocation of buffer and bandwidth

Relative cost of bandwidth versus buffer

Large deviations theory:

$$\text{Loss Probability} \approx \frac{1}{\sigma\theta\sqrt{2\pi N}} e^{-N(\sup_\theta[\theta(ct+b)-\log Ee^{\theta A(0,t))}])}$$

if both bandwidth and buffer are increased proportional to the number of sources!!



Morrison expression for loss as a function of allocated buffer, bandwidth, and number of on/off sources:

$$G(x) \approx \frac{1}{2}\sqrt{\frac{r}{f(\gamma)[\gamma+\lambda(1-\gamma)]N}}\exp\left(-N\,I(\gamma)\right)\exp\left(-2\sqrt{\{f(\gamma)[\gamma+\lambda(1-\gamma)]N\}}\right)\exp\left(-g(\gamma)x\right)$$

overflow prob.

γ — number of sources

γ, λ — bandwidth

buffer

Express as a Taylor series in terms of bandwidth above average & number of sources

Buffer vs. bandwidth at a fixed overflow probability:

$$x = c_1\delta^{-1}N^0 + c_2\delta^0 + c_3\delta^1 N^0$$

$c \quad c$

buffer

number of sources

bandwidth above average

Minimize total cost
if buffer and bandwidth have per unit costs

$$\delta = \sqrt{\frac{c_1}{m + c_3}} \frac{1}{\sqrt{N}} + O(\frac{1}{N^{\frac{3}{4}}})$$

bandwidth above average

ratio of cost of buffer to bandwidth

number of sources

Feasible choices of bandwidth and buffer per source

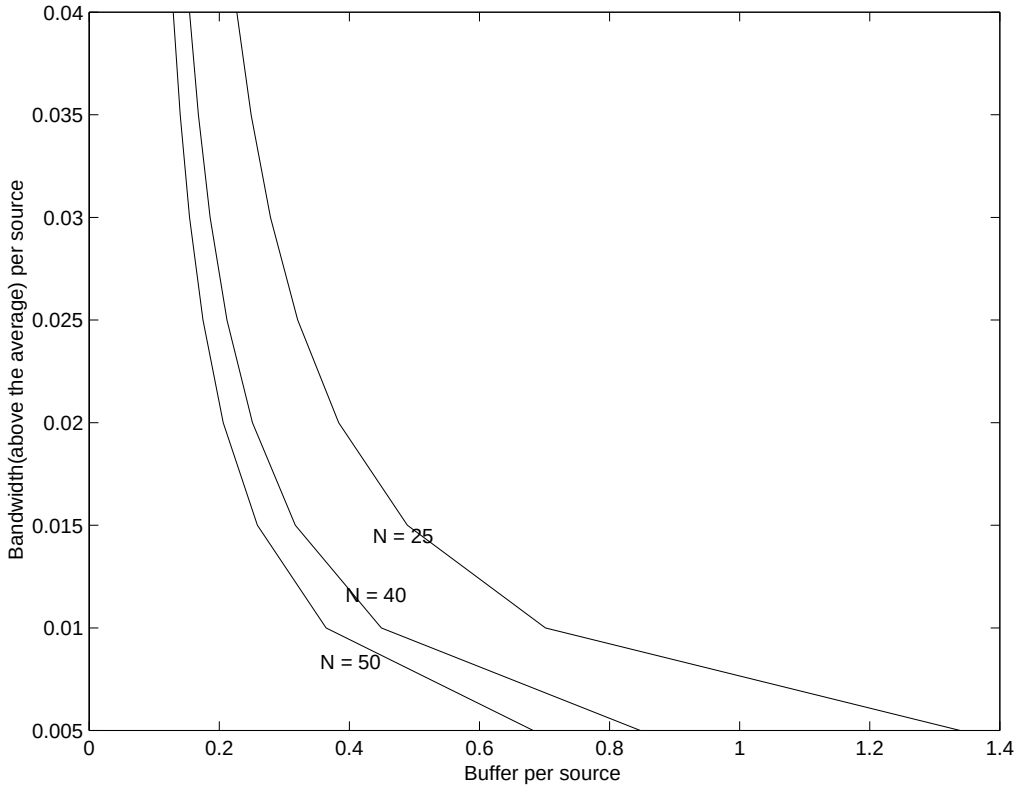


Figure 5.10: Variation of buffer, bandwidth above the average, δ, and number of sources N

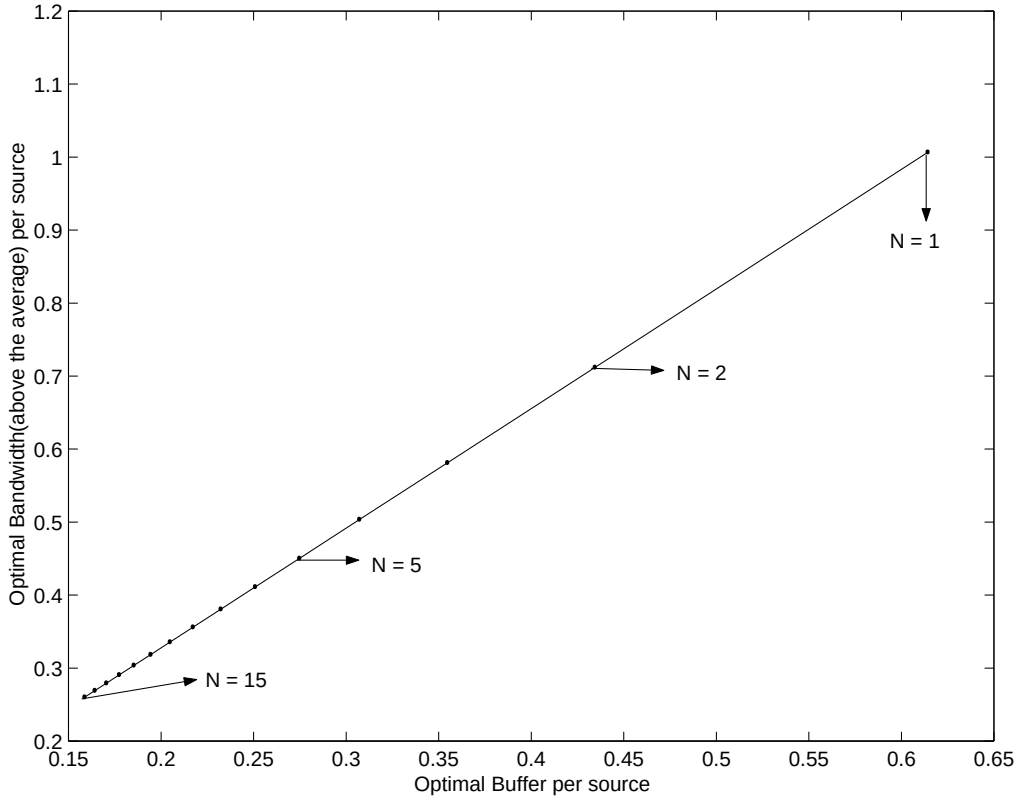
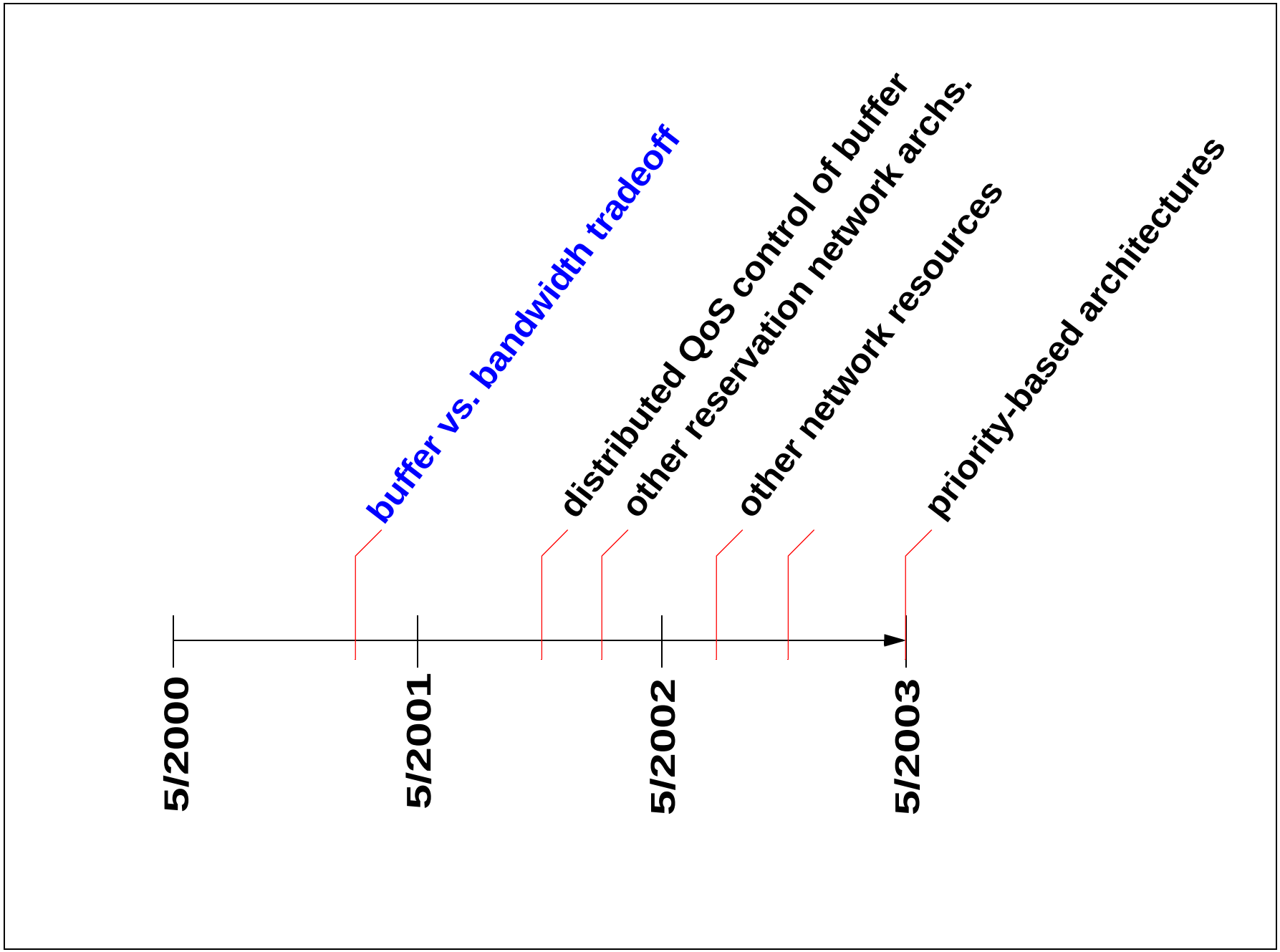


Figure 5.12: Variation of the *optimal resource allocation per source* with number of sources in the system



• *Build connections between characterizations of traffic flows, QoS requests, and network resource availability*

Traffic char.:

AT&T/Renesys?

Rice?

U. Maryland?

QoS Char.:

Network Res. Arch.:

Berkeley

UIUC?

Purdue?